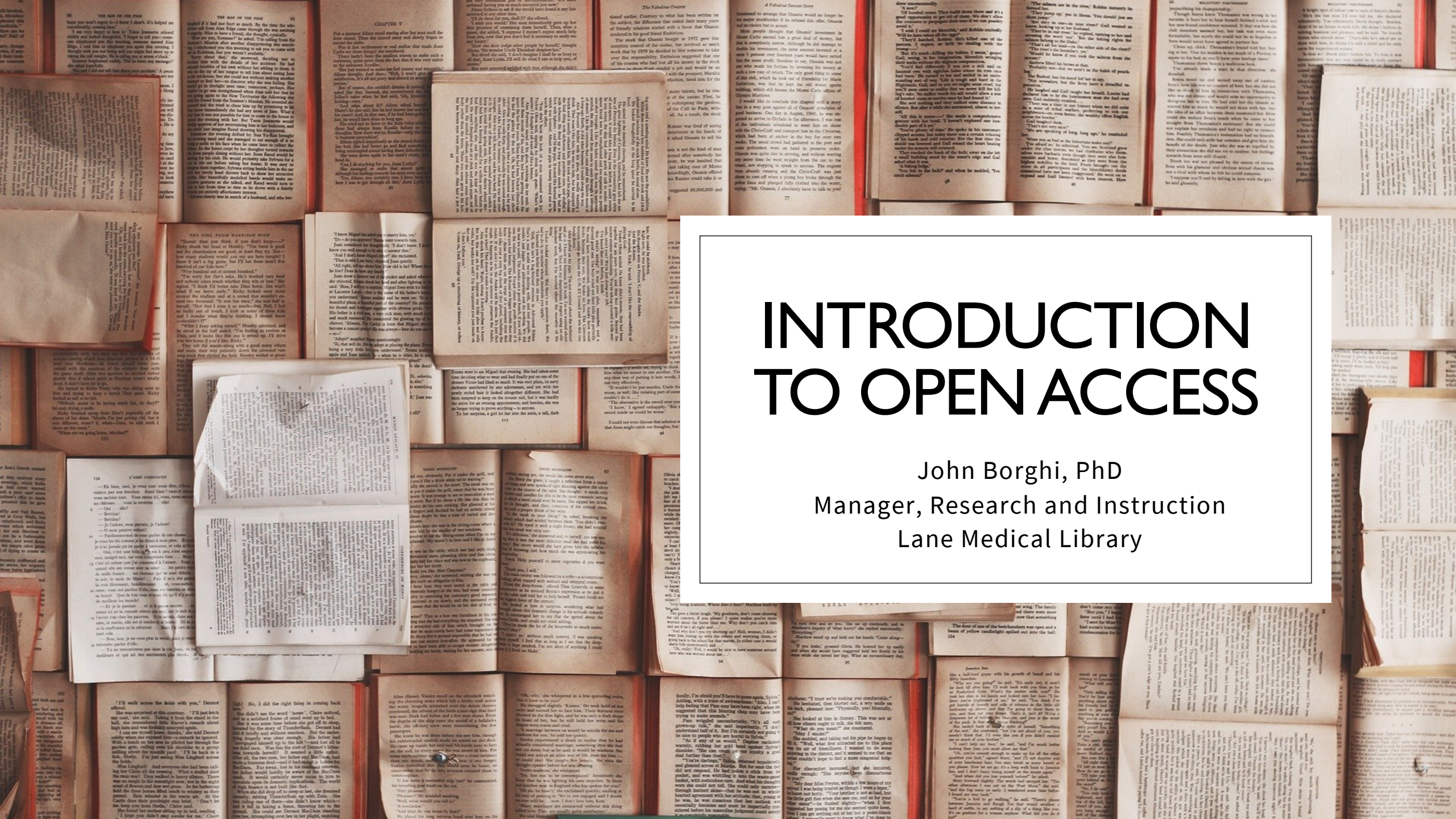


INTRODUCTION TO OPEN ACCESS

John Borghi, PhD
Manager, Research and Instruction
Lane Medical Library

OA Week at Lane Library

Enhancing Your Researcher Profile with ORCID | October 19, 2:00 pm

A Brief Introduction to Preprints
October 20, 2:00 pm

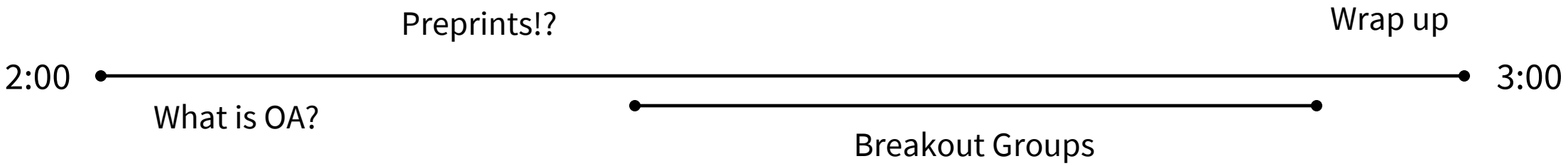
Introduction to Dryad (Data Sharing)
October 22, 2:00 pm

Open Science Reading Group
October 27, 2:00 pm

All events are free of charge and conducted virtually.

Visit Lane.Stanford.edu to register for our upcoming classes and events.

Today's Agenda



Reading for today:

Fraser, N., et al. (2020). Preprinting a pandemic: The role of preprints in the COVID-19 pandemic. *BioRxiv*. https://doi.org/10.1101/2020.05.22.111294

Additional reading:

Penfold, N. C., & Polka, J. K. (2020). Technical and social issues influencing the adoption of preprints in the life sciences. *PLOS Genetics*, *16*(4), e1008565. https://doi.org/10.1371/journal.pgen.1008565

Suber, P. (2012). What is Open Access? In *Open Access (pp 1-27). MIT Press.* https://archive.org/details/9780262517638OpenAccess/page/n13/mode/2up

What is Open Access?

By "**open access**" to [peer reviewed] literature, we mean its free availability on the public internet, permitting any users to read, download, copy, distribute, print, search, or link to the full texts of these articles, crawl them for indexing, pass them as data to software, or use them for any other lawful purpose, without financial, legal, or technical barriers other than those inseparable from gaining access to the internet itself.

Budapest Open Access Initiative (2002)

For a work to be **OA**, the copyright holder must consent in advance to let users “copy, use, distribute, transmit and display the work publicly and to make and distribute derivative works, in any digital medium for any responsible purpose, subject to proper attribution of authorship.”

Bethesda Statement on Open Access Publishing (2003)

Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities (2003)

What is Open Access?

“Open-access (OA) literature is digital, online, free of charge, and free of most copyright and licensing restrictions.”

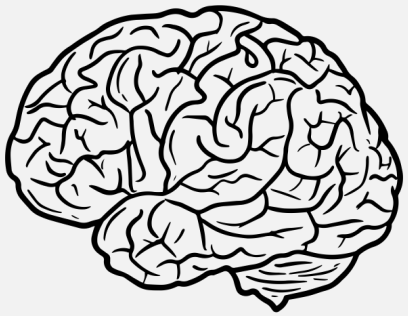
Suber (2004). A Brief Introduction to Open Access.

OA is a Continuum

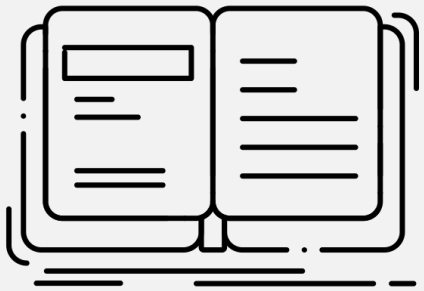
Access	Reader Rights	Reuse Rights	Copyrights	Author Posting Rights	Automatic Posting	Machine Readability	Access
OPEN ACCESS	Free readership rights to all articles immediately upon publication	Generous reuse & remixing rights (e.g., CC BY license)	Author holds copyright with no restrictions	Author may post any version to any repository or website	Journals make copies of articles automatically available in trusted third-party repositories (e.g., PubMed Central) immediately upon publication	Article full text, metadata, citations, & data, including supplementary data, provided in community machine-readable standard formats through a community standard API or protocol	OPEN ACCESS
	Free readership rights to all articles after an embargo of no more than 6 months	Reuse, remixing, & further building upon the work subject to certain restrictions & conditions (e.g., CC BY-NC & CC BY-SA licenses)	Author holds copyright, with some restrictions on author reuse of published version	Author may post final version of the peer-reviewed manuscript ("postprint") to any repository or website	Journals make copies of articles automatically available in trusted third-party repositories (e.g., PubMed Central) within 6 months	Article full text, metadata, citations, & data, including supplementary data, may be crawled or accessed through a community standard API or protocol	
	Free readership rights to all articles after an embargo greater than 6 months	Reuse (no remixing or further building upon the work) subject to certain restrictions and conditions (e.g., CC BY-ND license)	Publisher holds copyright, with some allowances for author and reader reuse of published version	Author may post final version of the peer-reviewed manuscript ("postprint") to certain repositories or websites	Journals make copies of articles automatically available in trusted third-party repositories (e.g., PubMed Central) within 12 months	Article full text, metadata, & citations may be crawled or accessed without special permission or registration	
	Free and immediate readership rights to some, but not all, articles (including "hybrid" models)	———	Publisher holds copyright, with some allowances for author reuse of published version	Author may post submitted version/draft of final work ("preprint") to certain repositories or websites	———	Article full text, metadata, & citations may be crawled or accessed with permission	
CLOSED ACCESS	Subscription, membership, pay-per-view, or other fees required to read all articles	No reuse rights beyond fair use/ limitations & exceptions to copyright (all rights reserved copyright) to read	Publisher holds copyright, with no author reuse of published version beyond fair use	Author may not deposit any versions to repositories or websites	No automatic posting in third-party repositories	Article full text & metadata not available in machine-readable format	CLOSED ACCESS

How Open Is It? – From SPARC, PLOS, and OASPA

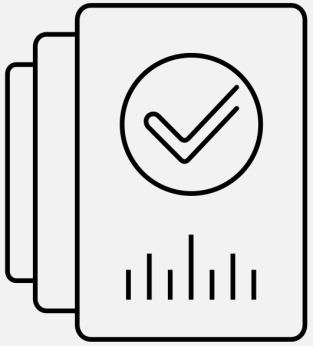
Traditional Scholarly Publishing



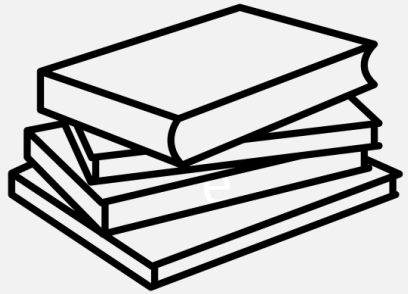
Research



Writing

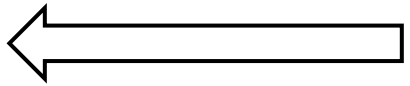


Evaluation



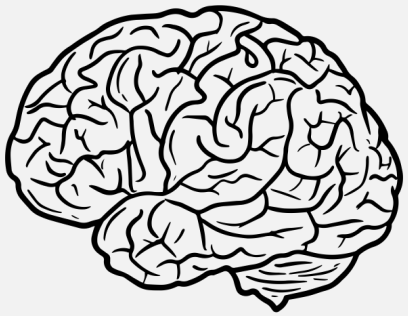
Publishing

Reading

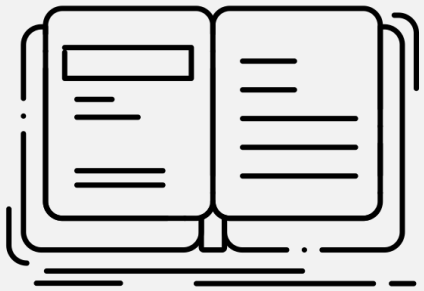


- Authors pay nothing (or comparatively little) to publish their work.
- Readers (libraries) pay fees to access journal articles.

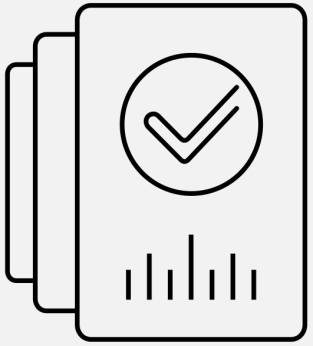
Open Access Publishing (Gold OA)



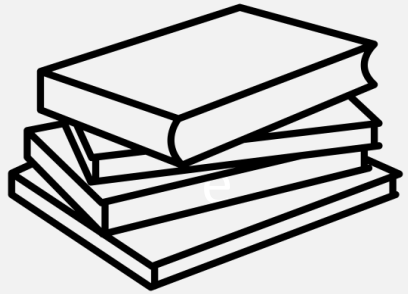
Research



Writing

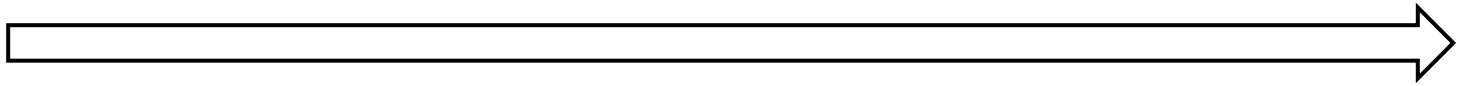


Evaluation



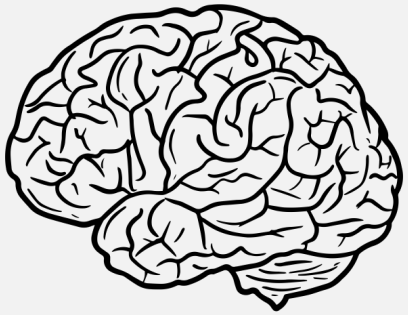
Publishing

Reading



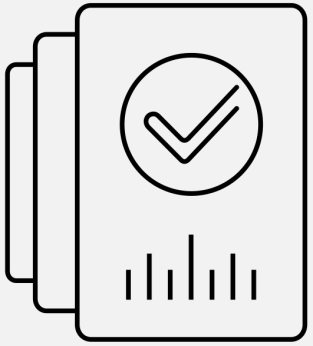
- Readers pay nothing to access journal articles.
- Authors often pay article processing charges to publish their work.

Hybrid Publishing and Bronze OA

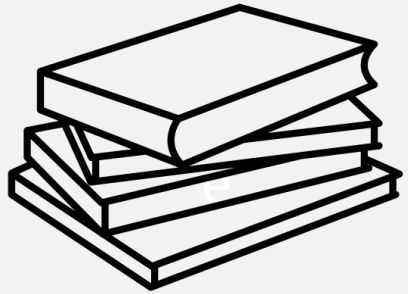


Research

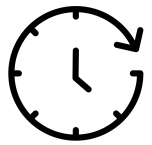
Writing



Evaluation



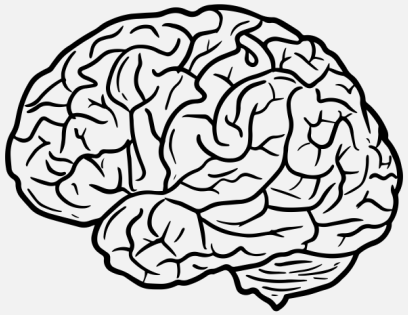
Publishing



Reading

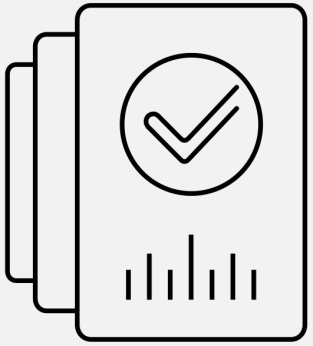
- Authors submit work to a publisher who typically charges readers for access.
- Publisher makes article free to read for (or after) a certain period or if authors pay a fee.

Self Archiving (Green OA)

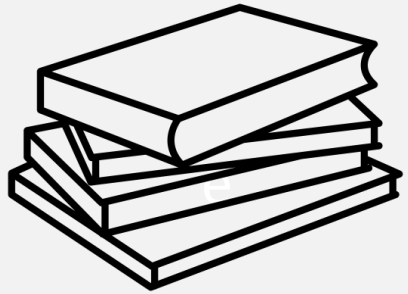


Research

Writing



Evaluation



Publishing

Reading

Authors submit articles for publication and deposit a version in an archive.

Archiving

Readers may pay to access publisher copy, can access archived copy for free.

What is a preprint?

Preprint – A version of an article that has not yet gone through the peer review process and does not include the journal's typesetting or formatting.

- **Post-print** - A version of an article that has gone through the peer review process and (probably) doesn't include the journal's typesetting or formatting. Also sometimes called an *author copy* or *author version*.
- **Publisher's Version** – The version of an article that appears on a journal website and/or in the print version. Also sometimes called the *version of record*.

Sharing Preprints Can…

- Help authors establish the priority of their work (prevent "scooping")
- Serve as "interim evidence of productivity"
- Accelerate scholarly communication and (potentially) scientific progress.
- Give authors feedback from a (potentially) broader group of reviewers.

A (Brief) History of Preprints

1961
NIH begins to circulate preprints through Information Exchange Groups.

1969
SLAC begins to distribute *Preprints in Particles and Fields*, a weekly lists of new preprints.

Ingelfinger rule announced at *New England Journal of Medicine*, drawing a line against prior publication.

1991
The LANL preprint archive is launched. Is later re-named arXiv and expanded beyond physics.

1996
Harold Varmus proposes e-BioMed, a platform to make biomedical research (including preprints) open access.

2013
bioRxiv is launched as the preprint server for biology.

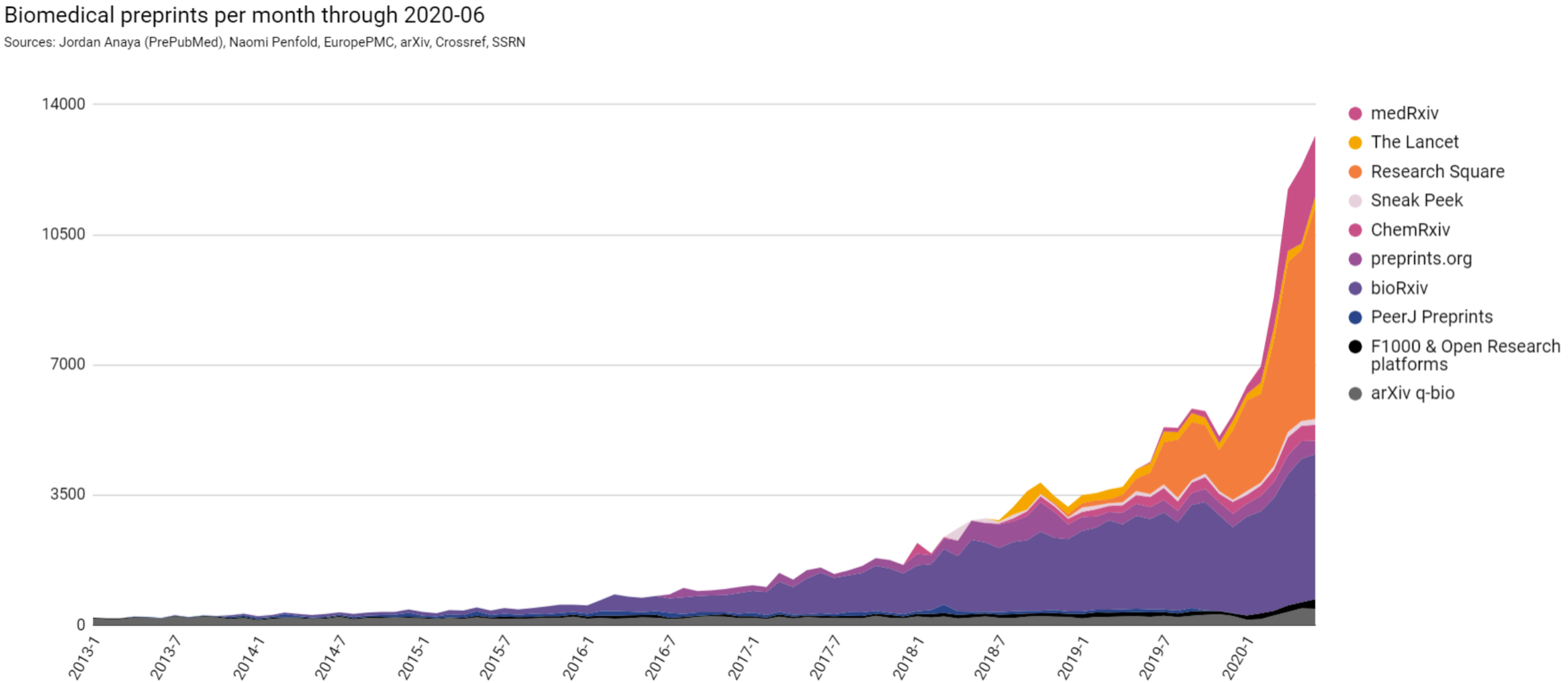
2019
MedRxiv is launched as the preprint server for the health sciences.

2020
While sheltering in place, everyone starts writing, submitting, reading, and talking about preprints.

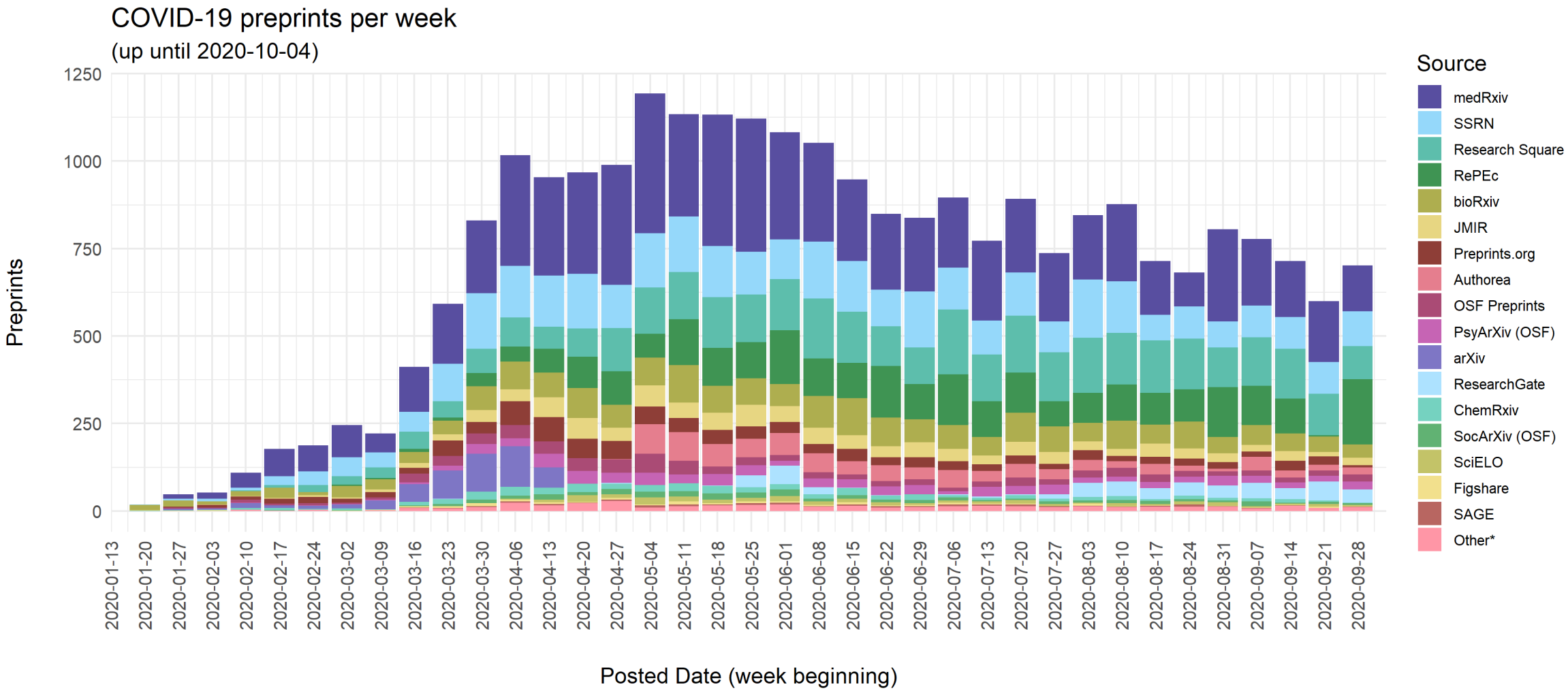
For more info, see Cobb, 2017

“Physicians have not been at the forefront of electronic communication, but many are beginning to use e-mail and to explore the Internet. Electronic communication is likely to become critically important in medicine, and it behooves physicians to become competent in using it. In the future, when the Internet is widely used by physicians, information about advances that may have an immediate effect on the health of individuals or populations may best be communicated through this medium. **But medicine is not physics: the wide circulation of unedited preprints in physics is unlikely to have an immediate effect on the public's well-being even if the material is biased or false. In medicine, such a practice could have unintended consequences that we all would regret**.”

Kassirer and Angell, 1995



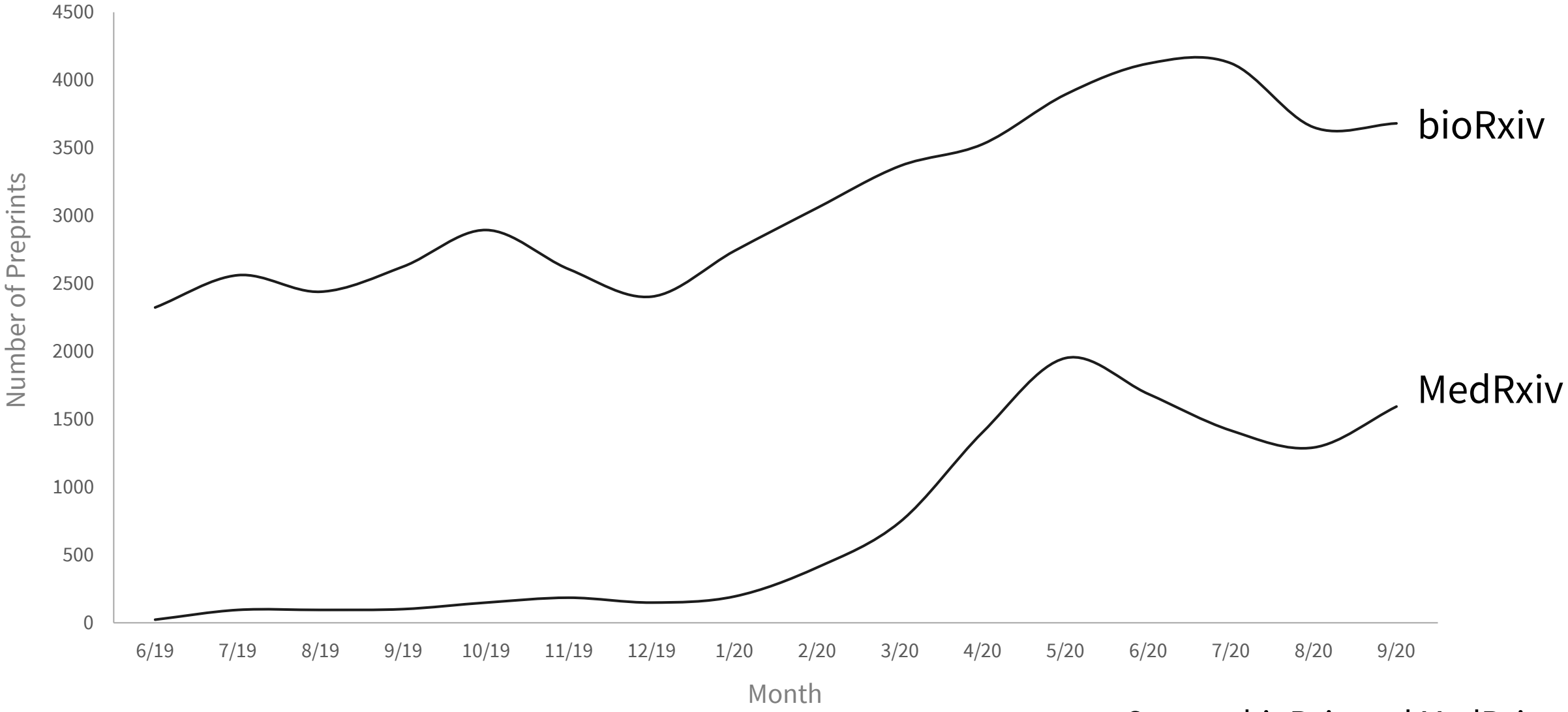
[Polka & Penfold, 2020]



* 'Other' refers to preprint repositories containing <50 total relevant preprints. These include: AfricArXiv (OSF), AgriXiv (OSF), BioHackrXiv (OSF), Cambridge University Press, Copernicus GmbH, EcoEvoRxiv (OSF), EdArXiv (OSF), engrXiv (OSF), ESSOAR, Frenxiv (OSF), INA-Rxiv (OSF), IndiaRxiv (OSF), LawArXiv (OSF), MediArXiv (OSF), MetaArXiV (OSF), NutriXiv (OSF), ScienceOpen, SportRxiv (OSF), Techrxiv (IEEE), WHO, Zenodo.

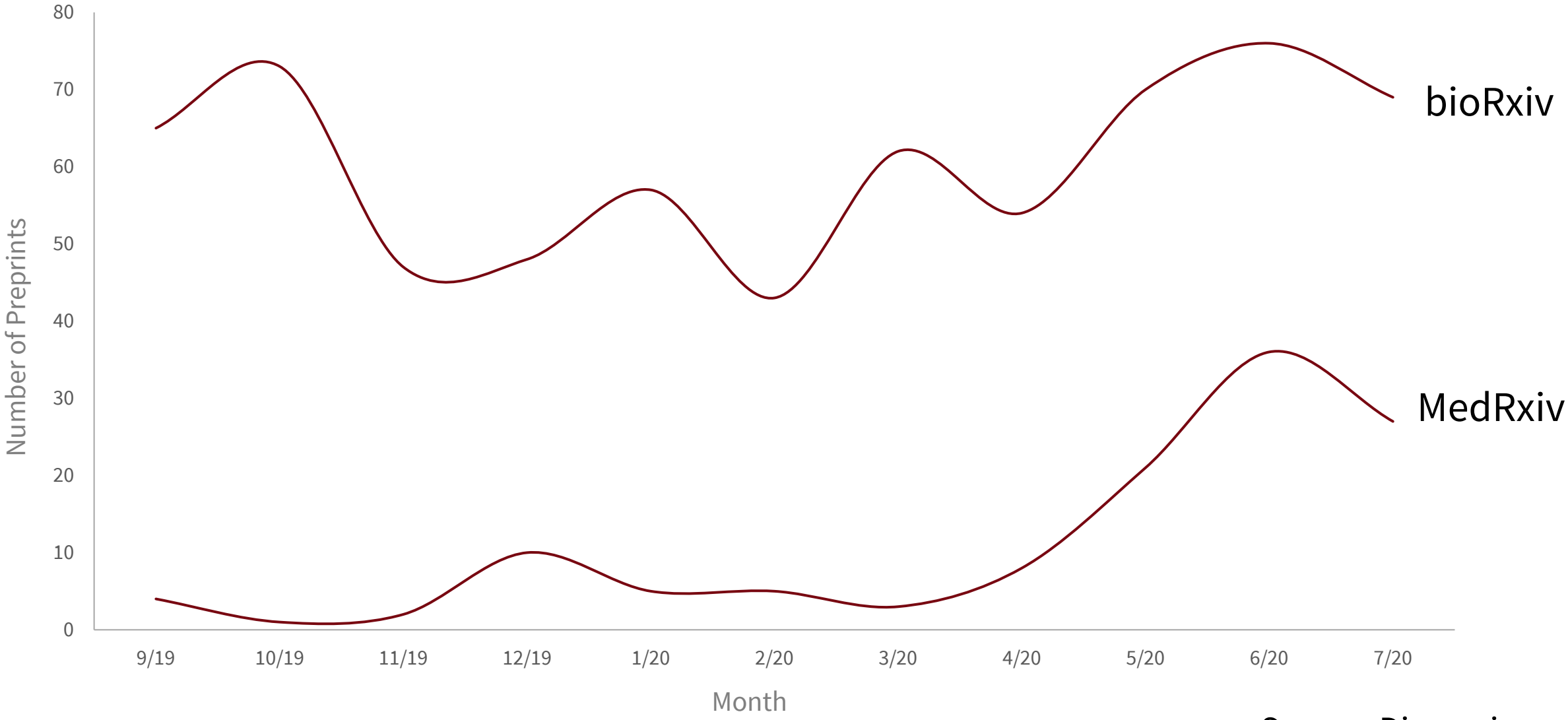
Fraser and Kramer, 2020

Preprints added to bioRxiv and MedRXiv



Source: bioRxiv and MedRxiv

Preprints posted by Stanford authors



Source: Dimensions

Table 1. Top 10 cited COVID-19 preprints

Rank	Source	doi	Title	Posted date	Citations
1	biorxiv	10.1101/2020.02.07.937862	Severe acute respiratory syndrome-related coronavirus - The species and its viruses, a statement of the Coronavirus Study Group	11/02/2020	127
2	medrxiv	10.1101/2020.02.06.20020974	Clinical characteristics of 2019 novel coronavirus infection in China	09/02/2020	126
3	medrxiv	10.1101/2020.01.23.20018549	Novel coronavirus 2019-nCoV: early estimation of epidemiological parameters and epidemic predictions	24/01/2020	112
4	biorxiv	10.1101/2020.01.22.914952	Discovery of a novel coronavirus associated with the recent pneumonia outbreak in humans and its potential bat origin	23/01/2020	93
5	biorxiv	10.1101/2020.01.26.919985	Single-cell RNA expression profiling of ACE2, the putative receptor of Wuhan 2019-nCov	26/01/2020	83
6	biorxiv	10.1101/2020.01.31.929042	The novel coronavirus 2019 (2019-nCoV) uses the SARS-coronavirus receptor ACE2 and the cellular protease TMPRSS2 for entry into target cells	31/01/2020	79
7	biorxiv	10.1101/2020.01.30.927806	The digestive system is a potential route of 2019-nCov infection: a bioinformatics analysis based on single-cell transcriptomes	31/01/2020	74
8	medrxiv	10.1101/2020.02.10.20021675	Epidemiological and clinical features of the 2019 novel coronavirus outbreak in China	11/02/2020	62
9	biorxiv	10.1101/2020.02.03.931766	Specific ACE2 Expression in Cholangiocytes May Cause Liver Damage After 2019-nCoV Infection	04/02/2020	49
10	medrxiv	10.1101/2020.03.03.20028423	Epidemiology and Transmission of COVID-19 in Shenzhen China: Analysis of 391 cases and 1,286 of their close contacts	04/03/2020	48

Fraser et al. 2020

Table 2. Top 10 tweeted COVID-19 preprints

Rank	Source	doi	Title	Posted date	Tweets	News articles	Blogs
1	medrxiv	10.1101/2020.04.14.20062463	COVID-19 Antibody Seroprevalence in Santa Clara County, California	17/04/2020	29984	328	24
2	biorxiv	10.1101/2020.01.30.927871	Uncanny similarity of unique inserts in the 2019-nCoV spike protein to HIV-1 gp120 and Gag	31/01/2020	18587	92	17
3	medrxiv	10.1101/2020.04.04.20053058	Indoor transmission of SARS-CoV-2	07/04/2020	17494	67	9
4	medrxiv	10.1101/2020.03.22.20040758	Efficacy of hydroxychloroquine in patients with COVID-19: results of a randomized clinical trial	30/03/2020	15337	117	15
5	medrxiv	10.1101/2020.03.09.20033217	Aerosol and surface stability of HCoV-19 (SARS-CoV-2) compared to SARS-CoV-1	10/03/2020	13407	333	27
6	biorxiv	10.1101/2020.03.13.990226	Reinfection could not occur in SARS-CoV-2 infected rhesus macaques	14/03/2020	10870	225	19
7	medrxiv	10.1101/2020.04.16.20065920	Outcomes of hydroxychloroquine usage in United States veterans hospitalized with Covid-19	21/04/2020	10512	329	15
8	medrxiv	10.1101/2020.03.30.20048165	Association of BCG vaccination policy with prevalence and mortality of COVID-19	06/04/2020	10435	3	0
9	medrxiv	10.1101/2020.03.17.20037713	A serological assay to detect SARS-CoV-2 seroconversion in humans	18/03/2020	8094	153	13
10	medrxiv	10.1101/2020.03.24.20042937	Correlation between universal BCG vaccination policy and reduced morbidity and mortality for COVID-19: an epidemiological study	28/03/2020	7427	77	5

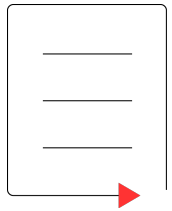
Fraser et al. 2020

Reviewing Preprints

BioRxiv – “All articles are screened on submission for offensive, dangerous, and/or non-scientific content and are checked for plagiarism.”
MedRxiv – “All manuscripts are screened on submission for plagiarism, non-scientific content, inappropriate article types, and material that could potentially endanger the health of individual patients or the public. The latter may include, but is not limited to, studies describing dual-use research and work that challenges or could compromise accepted public health measures and advice regarding infectious disease transmission, immunization, and therapy.”

There are a variety of tools to see/write reviews of preprints:

PUBPEER
The online journal club



PREREVIEW

publons

Let’s talk about OA and preprints!

Please keep in mind

Open Science is a broad umbrella (also, a buffet).

Let’s encourage understanding, participation, and inclusion.

- Nobody can do everything.
- Motivations differ.
- It’s not all or nothing.
- We are all learning.

Adapted from Mozilla Open Leadership Framework, Whitaker & Guest (2020)

Welcome Back!

- Anybody have anything from their discussion they want to share with the broader group?
- What (other) topics or issues would you like to discuss with the broader group?

Next Time

- Next meeting is November 23rd at 2:00 | Same Zoom information
- The topic is “Introduction to Data Sharing”
- Please suggest readings, discussion prompts, etc!

Some data (as of Oct 26th, 2020)

Of the 70,848 COVID-19-related papers/preprints in PubMed Central, only 7,197 **(10.16%)*** have a data citation or data availability statement attached.